

Optical Music Recognition in der Bayerischen Staatsbibliothek

Zusammenfassung: Die Bayerische Staatsbibliothek hat im Sommer 2016 ein Projekt zur automatischen Notenerkennung (Optical Music Recognition, OMR) gestartet. Dieser Beitrag beschreibt zunächst die Funktionsweise und die Qualität von OMR-Programmen und geht dann auf die bisherigen Erfahrungen der Bayerischen Staatsbibliothek mit OMR ein. Anschließend werden die Anwendungsszenarien skizziert, die die Bayerische Staatsbibliothek aufsetzend auf den OMR-Daten entwickeln wird.¹

Schlüsselwörter: Optical Music Recognition, Notenerkennung, Melodie-Suche

Optical Music Recognition in the Bavarian State Library

Abstract: In the summer of 2016, the Bavarian State Library has started a project on optical music recognition (OMR). This paper first describes the functionality and quality of OMR-programs and then depicts the OMR-experiences of the Bavarian State Library. Finally, the use case scenarios are outlined that the Bavarian State Library will implement on top of the OMR data.

Key words: Optical music recognition, melody search

1 Einleitung

Die Volltext-Suche in digitalisierten Büchern ist schon sehr weit verbreitet. Der weltweit umfassendste Index für Volltextbücher wird von der Firma Google in ihrem Programm „Google Book Search“ angeboten. Aber auch viele Bibliotheken ermöglichen dem Benutzer inzwischen, in ihren Bibliothekskatalogen nicht nur in den Metadaten der Bestände, sondern auch direkt in den Inhalten zu suchen. Dazu werden die digitalisierten Buchseiten mit einem Texterkennungs-Programm bearbeitet („Optical Character Recognition“, OCR), das die Buchstaben und Wörter in den Digitalisaten identifiziert. Man spricht hier von der sogenannten „inhaltsbasierten Suche“. Die inhaltsbasierte Suche in Musikdrucken steckt dagegen noch in den Anfängen. Zwar gibt es analog zu den OCR-Programmen schon seit einigen Jahren Programme zur automatischen Erkennung von

¹ Dieser Beitrag basiert auf dem von Jürgen Diet und Sanu Pulimootil verfassten Artikel in der Zeitschrift „Bibliotheksmagazin“ (1) (2017) 38–40, mit dem Titel „Bayerische Staatsbibliothek startet Projekt zur automatischen Erkennung von Musiknoten“.

Musiknoten („Optical Music Recognition“, OMR), deren Fehlerquote im Vergleich zu den OCR-Programmen aber noch relativ hoch ist; zudem gibt es noch kaum Anwendungen, die eine inhaltsbasierte Suche in Musiknoten anbieten.

Im Rahmen ihrer umfassenden Digitalisierungsstrategie startete die Bayerische Staatsbibliothek im Sommer 2016 ein Projekt im Bereich Optical Music Recognition. Als eine der ersten Einrichtungen weltweit plant die Bayerische Staatsbibliothek den institutionellen Einsatz von OMR-Programmen als Basis für eine inhaltsbasierte Suche in historischen Musikquellen. In einer ersten Phase wird getestet, welche OMR-Programme für die digitalisierten Notendrucke der Bayerischen Staatsbibliothek gute Ergebnisse liefern. Anschließend konzipiert und entwickelt das Projekt-Team Anwendungen, die auf den OMR-Ergebnissen aufsetzen, z.B. eine Suche in digitalisierten Notendruckten nach einer vorgegebenen Melodie.

2 Was kann Optical Music Recognition heute?

Die automatische Erkennung von Musiknoten (Optical Music Recognition) ist eine größere Herausforderung als die automatische Texterkennung (Optical Character Recognition), da ein Notentext in der Regel sehr viel komplizierter ist als ein gewöhnlicher Text. Beim OCR geht es darum, Buchstaben aus einer überschaubaren Menge eines Alphabets zu erkennen und sie zu Wörtern und Sätzen zusammenzufassen. Die Menge an zu erkennenden Zeichen in einem Notentext ist dagegen sehr viel größer und ihr Zusammenwirken deutlich komplexer.

Die Entwicklung von OMR-Software wird zudem dadurch erschwert, dass für sie im Vergleich zu OCR-Programmen nur ein kleiner Markt existiert. Von daher wurde bisher sowohl von Firmen wie auch von öffentlicher Seite deutlich mehr Geld und Aufwand in die Erstellung und Weiterentwicklung von OCR-Programmen gesteckt als in OMR-Software. Trotz dieser Ausgangssituation gibt es schon eine Reihe von kommerziellen und Open-Source-OCR-Programmen.²

Zwei Zeitschriften- bzw. Konferenzbeiträge aus den Jahren 2012³ und 2015⁴ geben einen Überblick über den derzeitigen Stand bei OMR sowie die speziellen Herausforderungen beim Erkennen von Musiknoten. Es ist derzeit sehr schwierig, die Qualität von OMR-Programmen zu beurteilen und Vergleiche zwischen verschiedenen OMR-Programmen anzustellen. Im Textbereich lässt sich die Qualität von OCR-Programmen vergleichsweise einfach darstellen. Mit einer Erkennungsrate in

² Im Wikipedia-Eintrag „Optische Notenerkennung“ werden 14 OMR-Programme aufgelistet, siehe https://de.wikipedia.org/wiki/Optische_Notenerkennung (Zugriff am 2.8.2017).

³ Rebelo, Ana; Fujinaga, Ichiro; Paszkiewicz, Filipe; Marcal, Andre R.S.; Guedes, Carlos; Cardoso, Jaime S.: Optical music recognition: state-of-the-art and open issues, *International Journal of Multimedia Information Retrieval* (1) (2012) 173–90, <http://dx.doi.org/10.1007/s13735-012-0004-6> (Zugriff am 4.8.2017).

⁴ Novotný, Jiří; Pokorný, Jaroslav: Introduction to Optical Music Recognition: Overview and Practical Challenges, *International Workshop on Databases, Texts, Specifications, Objects 2015 (DATESO 2015)*, Czech Republic, April 14, 2015, ed. By Martin Necasky, Jaroslav Pokorný, Pavel Moravec, <http://ceur-ws.org/Vol-1343/paper6.pdf> (Zugriff am 4.8.2017).

Prozent kann man angeben, welcher Prozentsatz der Zeichen bzw. der Wörter eines Textes fehlerfrei identifiziert wurde. Aufgrund der hohen Komplexität eines Notendrucks lässt sich diese Vorgehensweise nicht so leicht auf OMR-Programme übertragen. Don Byrd und Jakob Grue Simonsen beklagten im Jahr 2015 diesen Zustand in ihrem Artikel im *Journal of New Music Research* und schlugen eine Standard-Testumgebung für OMR vor („Standard Testbed for Optical Music Recognition“).⁵ Obwohl ein Jahr später weitere Vorschläge für eine derartige OMR-Testumgebung gemacht wurden,⁶ gibt es bisher kein System für die standardisierte Evaluierung von OMR-Programmen. Eine geeignete Institution für diese Aufgabe wäre die International Society for Music Information Retrieval (ISMIR),⁷ die auf ihren jährlichen Tagungen im Rahmen von MIREX (Music Information Retrieval Evaluation Exchange)⁸ verschiedene standardisierte Tests durchführt, um unterschiedliche Algorithmen und Programme im Bereich des Music Information Retrievals zu testen und zu vergleichen.

Die nächsten drei Abbildungen zeigen Beispiele für OMR-Ergebnisse. Das kurze Notenbeispiel aus Abb. 1 wird mit zwei verschiedenen OMR-Programmen erkannt. Beim Ergebnis des OMR-Programms A in Abb. 2 zeigen sich Fehler bei der Satzbezeichnung („Allegro maestoso“), beim Rhythmus im 7. Takt und bei der letzten Note. Das OMR-Programm B lieferte ein besseres Ergebnis und hat nur die Dynamik-Bezeichnung zu Beginn („ff“) nicht erkannt (siehe Abb. 3).



Abb. 1: Beginn der 2. Fagott-Stimme von Mottl, Felix / Gluck, Christoph Willibald: Ballet-Suite, Balletstücke aus Opern von Chr. W. von Gluck; 4 Mus.pr. 62427

⁵ Byrd, Don und Simonsen, Jakob Grue: Towards a Standard Testbed for Optical Music Recognition: Definitions, Metrics, and Page Images, *Journal of New Music Research* 44(3) (2015) 169–95, <http://dx.doi.org/10.1080/09298215.2015.1045424> (Zugriff am 3.8.2017).

⁶ Hajič jr., Jan; Novotný, Jiří; Pecina, Pavel; Pokorný, Jaroslav: Further Steps Towards a Standard Testbed for Optical Music Recognition, 17th International Society for Music Information Retrieval Conference 2016, https://wp.nyu.edu/ismir2016/wp-content/uploads/sites/2294/2016/07/289_Paper.pdf (Zugriff am 3.8.2017).

⁷ www.ismir.net (Zugriff am 3.8.2017).

⁸ www.mirex.org (Zugriff am 3.8.2017).



Abb. 2: Ergebnis des OMR-Programms A



Abb. 3: Ergebnis des OMR-Programms B

Eine weitere Veranschaulichung von OMR-Ergebnissen findet man auf einer von Craig Sapp erstellten Webseite,⁹ auf der er die Ergebnisse der OMR-Programme SmartScore und SharpEye für die ersten 20 Takte von Ludwig van Beethovens Klaviersonate op. 2, Nr. 1 Takt für Takt erläutert.

Bei der Qualität der OMR-Ergebnisse spielen viele Faktoren eine Rolle, u.a. die Komplexität der Musik, das Jahr des Notendrucks und die Bildqualität. Bei einigen Scans von Notendrucken sind die Notenzeilen nicht ganz horizontal, sondern leicht gebogen, was für die OMR-Programme eine besondere Herausforderung darstellt. Ein besonderes Augenmerk muss auf die Erkennung der musikalischen Symbole zu Beginn des Musikstückes und zu Beginn jeder Notenzeile gelegt werden. Wenn der Notenschlüssel, die Vorzeichen oder die Taktart nicht richtig erkannt werden, dann führt dies zu Folgefehlern in jedem weiteren Takt. Zur Verbesserung der OMR-Ergebnisse kann es sinnvoll sein, verschiedene Vorverarbeitungsschritte vor dem Aufruf des OMR-Programmes einzufügen, z.B. die Änderung des Bildformates oder das Entfernen von Bildrändern.

Es gibt bisher nur wenige Anwendungen, bei denen OMR im großen Stil angewendet wird. Meist werden die OMR-Daten dabei für eine inhaltsbasierte Suche verwendet, also einer Suche nicht nach Metadaten (Komponist, Werktitel, Besetzung, Verlag, u.a.), sondern nach dem Inhalt der Musik analog zu einer Volltext-Suche bei textuellen Medien. Eine dieser Anwendungen ist die von Vladimir Viros Firma Peachnote entwickelte Melodie-Suche, für die er die Notendigitalisate der Petrucci Library und weiterer Musikbibliotheken mit OMR bearbeitet hat.¹⁰

⁹ <https://ccrma.stanford.edu/~craig/mro-compare-beethoven> (Zugriff am 3.8.2017).

¹⁰ Viro, Vladimir: Peachnote: Music Score Search and Analysis Platform, 12th International Society for Music Information Retrieval Conference 2011, <http://ismir2011.ismir.net/papers/PS3-1.pdf>, (Zugriff am 8.8.2017).

Music Ngram Viewer



Please enter a melody or a sequence of chords (advanced use)



62 0 1 2 0 -2 -1 -2 -2

melody

Search

Petrucchi Music Library

Smoothing: 0

☐ Normalized

Keyboard shortcuts

filter search results (e.g. Mozart, winds, or quartet)

- Solos for Treble Instrument, Especially Soprano Recorder (Kimberling, Clark) [YouTube](#) [\[score\]](#) page [100](#)
Kimberling, Clark (2000) [\[score\]](#) pages [11,34, 107,149](#)
- Il barbiere di Siviglia (Rossini, Gioacchino) [YouTube](#) [\[score\]](#) page [101](#)
Rossini, Gioacchino (1816)
- The Free-Lance (march) (Sousa, John Philip) [YouTube](#) [\[score\]](#) pages [3,4, 5,13](#)
Sousa, John Philip (1906)
- School of Intonation on an Harmonic Basis for Violin, Op.11 (Ševčík, Otakar) [YouTube](#) [\[score\]](#) page [14](#)
Ševčík, Otakar (1922)
- Symphony No.9, Op.125 (Beethoven, Ludwig van) [YouTube](#) [\[score\]](#) pages [14,30](#)
Beethoven, Ludwig van (1822)
- 6 String Quartets, Op.8 (Giordani, Tommaso) [YouTube](#) [\[score\]](#) pages [9,19](#)
Giordani, Tommaso (1775)
- 4 Trio Sonatas, WK 98-101 (Abel, Carl Friedrich) [YouTube](#) [\[score\]](#) pages [3,10](#)
Abel, Carl Friedrich (1785)

[next](#)

you can also browse using the chart

Search for scores:

1508 - 1766	1767 - 1790	1791 - 1825	1826 - 1887	1888 - 2012	62 0 1 2 0 -2 -1 -2 -2
-----------------------------	-----------------------------	-----------------------------	-----------------------------	-----------------------------	------------------------

Abb. 4: Startseite der Peachnote-Suche (www.peachnote.com, Zugriff am 8.8.2017)

3 Erfahrungen der Bayerischen Staatsbibliothek mit Optical Music Recognition

Die Bayerische Staatsbibliothek machte erste Erfahrungen mit OMR im PROBADO-Projekt,¹¹ das von 2006 bis 2011 von der Deutschen Forschungsgemeinschaft gefördert wurde. In diesem Projekt wurde untersucht, wie multimediale Objekte in bibliothekarische Geschäftsgänge integriert werden können.¹² Als Anwendungsfälle wurden 3D-Architekturmodelle und Musik in Form von Notendruckern und Audio-Aufnahmen verwendet. Die Bayerische Staatsbibliothek war zusammen mit dem Projektpartner Universität Bonn für den Musik-Bereich bei PROBADO verantwortlich. Es entstand eine prototypische Anwendung, die die synchronisierte Wiedergabe von Audio-Aufnahmen und dazugehörigen Notendruckern ermöglichte. Voraussetzung hierfür war die automatisierte inhaltliche Erschließung sowohl der Audio-Aufnahmen mithilfe von Signalverarbeitungs-Verfahren

¹¹ www.probado.de (Zugriff am 7.8.2017).

¹² Berndt, René; et al: The PROBADO Project – Approach and Lessons Learned in Building a Digital Library System for Heterogeneous Non-textual Documents. In: Lalmas M., Jose J., Rauber A., Sebastiani F., Frommholz I. (eds) Research and Advanced Technology for Digital Libraries. ECDL 2010. Lecture Notes in Computer Science, vol 6273. Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-15464-5_37 (Zugriff am 7.8.2017).

wie auch der Notendrucke durch den Einsatz von OMR. Im PROBADO-Projekt wurde das kommerzielle OMR-Programm SharpEye in der Version V2.68 eingesetzt. Leider war es aus verschiedenen Gründen nicht möglich, die prototypische PROBADO-Musik-Anwendung in einen produktiven Betrieb an der Bayerischen Staatsbibliothek zu überführen. Ein Demo-Video, das mit dem PROBADO-Musik-Prototyp erstellt wurde, vermittelt jedoch einen Eindruck von den Projektergebnissen.¹³

Im Jahr 2014 startete an der McGill-University in Montreal das auf sieben Jahre angelegte und von Prof. Ichiro Fujinaga geleitete SIMSSA-Projekt („Single Interface for Music Score Searching and Analysis“¹⁴), in dem OMR-Verfahren und Anwendungen für die inhaltsbasierte Suche in Notendigitalisaten entwickelt werden. Die Bayerische Staatsbibliothek ist Partner im SIMSSA-Projekt und stellt ihre Notendigitalisate für die SIMSSA-Anwendung zur Verfügung. Außerdem prüft sie, ob die im SIMSSA-Projekt entwickelten Technologien in ihrem eigenen OMR-Projekt eingesetzt werden können.

In dem im Sommer 2016 gestarteten OMR-Projekt hat die Bayerische Staatsbibliothek zunächst Tests mit verschiedenen OMR-Programmen auf ausgewählten Notendigitalisaten aus ihrem Bestand durchgeführt (Audiveris,¹⁵ Capella Scan,¹⁶ SharpEye¹⁷ und SmartScore¹⁸). Audiveris ist ein open-source-Programm, bei den anderen drei OMR-Programmen handelt es sich um kommerzielle Software. Wie oben beschrieben, ist die Evaluierung von OMR-Programmen sehr schwierig. Die Testergebnisse sind daher zwangsläufig subjektiv und können nicht verallgemeinert werden. Es hat sich gezeigt, dass bei allen getesteten OMR-Programmen zahlreiche Erkennungsfehler auftreten. Die Bayerische Staatsbibliothek wird keine manuelle Korrektur der OMR-Daten vornehmen, da dies für die geplante Menge an OMR-Daten zu aufwändig wäre. Die Anwendungsszenarien, die mit den OMR-Daten konzipiert werden, müssen also fehlertolerant sein. Für den weiteren Verlauf ihres OMR-Projektes hat sich die Bayerische Staatsbibliothek für das OMR-Programm SmartScore entschieden. Dieses Programm wird nun dafür eingesetzt, um OMR-Daten für die von Carl Alexander herausgegebene 34-bändige Gesamtausgabe der Werke von Franz Liszt¹⁹ und für die von Friedrich

¹³ http://www.probado.de/demos/musik_interface_01/index.html (Zugriff am 4.8.2017).

¹⁴ <https://simssa.ca> und <http://www.mcgill.ca/newsroom/channels/news/mcgill-led-partnership-projects-receive-sshrc-funding-238500> (Zugriffe am 7.8.2017).

¹⁵ <https://github.com/audiveris> (Zugriff am 8.8.2017).

¹⁶ <http://www.capella-software.com/de/> (Zugriff am 8.8.2017).

¹⁷ <http://www.visiv.co.uk> (Zugriff am 8.8.2017).

¹⁸ <http://www.klemm-music.de/musitek/smartscore> (Zugriff am 8.8.2017).

¹⁹ <https://www.digitale-sammlungen.de/index.html?c=sammlung&projekt=1193812455&l=de> (Zugriff am 8.8.2017).

Chrysander herausgegebene 95-bändige Gesamtausgabe der Werke von Georg Friedrich Händel²⁰ zu erstellen.

4 Ausblick

Die in der Bayerischen Staatsbibliothek erzeugten OMR-Daten sollen im Laufe des Jahres 2018 dazu genutzt werden, um eine inhaltsbasierte Suche in den digitalisierten Musikdrucken zu ermöglichen. Dabei wird (wie bei der oben erwähnten Peachnote-Suche) nur der Tonhöhen-Verlauf berücksichtigt, alle anderen musikalischen Informationen (wie z.B. der Rhythmus und die musikalischen Vortragsbezeichnungen) werden außer Acht gelassen. Die Eingabe bei der Melodie-Suche wird über eine virtuelle Klaviatur erfolgen. Bei der Gestaltung der Melodie-Recherche kann die Bayerische Staatsbibliothek auf ihre im Zusammenhang mit der Implementierung des RISM-OPAC²¹ aufgebauten Erfahrungen zurückgreifen, der ebenfalls eine Melodie-Suche bietet. Das Internationale Quellenlexikon der Musik (*Répertoire International des Sources Musicales*, RISM) enthält mehr als eine Million Musikquellen aus aller Welt, die nach Autopsie erschlossen und manuell mit Musik-Incips versehen wurden.

Bei der inhaltlichen Erschließung von Musikquellen besteht heute noch ein großer Qualitätsunterschied zwischen den mit OMR-Programmen erzeugten Daten und manuell erstellten Daten (wie z.B. die Incips bei RISM). Die auf den OMR-Daten aufsetzenden Anwendungsszenarien müssen daher so robust implementiert werden, dass sie trotz der nicht fehlerfreien OMR-Daten noch brauchbare Ergebnisse liefern. Außerdem wird sich hoffentlich die Erkennungsrate bei OMR-Programmen in den nächsten Jahren mit neuen Programmversionen verbessern.



Jürgen Diet

Bayerische Staatsbibliothek

Ludwigstr. 16

D-80539 München

juergen.diet@bsb-muenchen.de

²⁰ <https://www.digitale-sammlungen.de/index.html?c=sammlung&projekt=1193214396&l=de> (Zugriff am 8.8.2017).

²¹ <https://opac.rism.info> (Zugriff am 4.8.2017).